The structure of triplex DNA in the gas phase

Annalisa Arcella[a], Guillem Portella[a] , Maria Luz Ruiz[b], Ramon Eritja[b] , Marta Vilaseca[c],

Valérie Gabelica,[d] and Modesto Orozco[a,e*]

[a] Joint IRB BSC research program in Computational Biology. Institute for Research in Biomedicine . Josep Samitier 1-5. Barcelona 08034. Spain

[b] Chemistry and Molecular Pharmacology Program. Institute for Research in Biomedicine. IQAC-CSIC, CIBER-BBN. Barcelona 08028. Spain.

[c] Mass Spectrometry Core Facility. Institute for Research in Biomedicine. Barcelona 08028. Spain.

[d] Department of Chemistry. University of Liège. Allée de la Chimie, Building B6c, B-4000 Liège, Belgium.

[e] Departament de Bioquímica i Biología Molecular. Facultat de Biología. Universitat de Barcelona. Avgda Diagonal 645. Barcelona 08028. Spain.

* Correspondence to Prof. Modesto Orozco: modesto.orozco@irbbarcelona.org

Supporting information:

Table S1. Summary of simulations performed for 12-mer (charge state -6) and 18-mer triplexes (charge states -7 and -8) in gas phase and aqueous solution for different charge states. See main text for details.

Triplex length	Environment	# Replicates	Temperature	Length (μs)	Simulation time (μs)
12-mer	Gas	10	300	1.0	10
	Gas	1	300	30	30
	Gas	1	300	1.0 x 5	5
	Gas	10	372	1.0	10
	Solution	1	300	0.1	0.1
18-mer	Gas	20 (10 replicates for each charge state: -7, -8)	300	1.0	20
	Gas	20 (10 replicates for each charge state: -7, -8)	372	1.0	20
	Solution	1	300	0.1	0.1
	Solution	1	373	0.1	0.1

Table S2: Summary of experimental conditions used on the two Synapt G1 HDMS mass spectrometers.

	Barcelona	**Liège**
Samples	pH = 7, 150 µM triplex, 5% isopropanol – 95% H_2O	pH = 3, 10 µM (12-mer) or 15 µM (18-mer) triplex, 100% H_2O
-Ionization source	Triversa NanoMate (Advion) chip-based nanoelectrospray source. Spray voltage = 1.65 kV, delivery pressure = 0.9 psi.	Standard electrospray source (Waters, Manchester, UK). Spray voltage = 2.2 kV, flow rate = 4 µL/min
Source pressure (pirani reading)	5.49 mbar	3.10 mbar
Cone voltage	40 V	30 V
Extraction cone	5 V	4 V
Source temperature	20 °C	25 °C
Desolvation temperature	not applicable	40 °C
p_{Ar} *in trap and transfer (pirani reading)*	4.67×10^{-2} mbar	2.59×10^{-2} mbar
Trap collision energy	10 V	4 V
Transfer collision energy	10 V	4 V
IMS cell bias	20 V	10 V, 15 V, 20 V, 25 V, and 30 V
p_{N2} *in IMS cell (pirani reading)*	0.546 ± 0.002 mbar	0.532 ± 0.001 mbar
Transfer wave velocity	248 m/s	248 m/s

Table S3. Oligonucleotides used as calibrant ions and their published collision cross sections (Ω).

Oligonucleotide name	Charge	m/z	Ω in Å^2	Reference[(a)]	Used in Liège	Used in Barcelona
Single strand d(TGT)	-1	874	238	*JASMS,* ***2003****, 14, 161*		X
Single strand d(GTT)	-1	874	240	*JASMS,* ***2003****, 14, 161*		X
Single strand dT_{10}	-3	992	446	*JACS,* ***1997****, 119, 9051*	X	
Single strand dT_{10}	-4	744	537	*JACS,* ***1997****, 119, 9051*	X	
Single strand dT_{10}	-5	595	627	*JACS,* ***1997****, 119, 9051*	X	
Single strand dT_{10}	-6	496	641	*JACS,* ***1997****, 119, 9051*	X	
Duplex $[d(CG)_2]_2$+Na	-3	789	352	*JACS,* ***2004****, 126, 15132*		X
Duplex $[d(CG)_3]_2$	-3	1194	430	*JACS,* ***2004****, 126, 15132*		X
Duplex $[d(CG)_{11}]_2$	-11	1223	1404	*JASMS,* ***2007****, 18, 1188*		X
Duplex [dGACCTAGATTCGAATCTAGGTC]$_2$	-11	1224	1529	*JASMS,* ***2007****, 18, 1188*		X
Duplex [dCAGTCACGGATTCGAATCCGTGACTG]$_2$	-13	1226	1776	*JASMS,* ***2007****, 18, 1188*		X
Quadruplex $[(dTGGGGT)_4•(NH_4)_3]$	-5	1499	775	*IJMS,* ***2006****, 253, 206*	X	X
Quadruplex $dGGG(TTAGGG)_3$	-4	1662	688	*JACS*, **2007**, 129, 895	X	X
Quadruplex $dGGG(TTAGGG)_3$	-5	1330	718	*JACS*, **2007**, 129, 895	X	

(a) JACS = J. Am. Chem. Soc.: JASMS = J. Am. Soc. Mass Spectrom.; IJMS = Int. J. Mass Spectrom.

Table S3 (continued).

Oligonucleotide name	Charge	m/z	Ω in Å^2	Reference[(a)]	Used in Liège	Used in Barcelona
Quadruplex [dGAGGGTGGGGAGGGTGGGGAAG•$(NH_4)_2$]	-4	1776	701	*JACS*, **2007**, 129, 895	X	
Quadruplex [dGAGGGTGGGGAGGGTGGGGAAG•$(NH_4)_2$]	-5	1419	757	*JACS*, **2007**, 129, 895	X	
Denatured quadruplex [dGAGGGTGGGGAGGGTGGGGAAG]	-5	1412	696	*JACS*, **2007**, 129, 895	X	
Quadruplex [dGAGGGTGGGGAGGGTGGGGAAG•$(NH_4)_2$]	-6	1182	807	*JACS*, **2007**, 129, 895	X	
Denatured quadruplex [dGAGGGTGGGGAGGGTGGGGAAG]	-6	1176	900	*JACS*, **2007**, 129, 895	X	
Quadruplex dimer [(dGAGGGTGGGGAGGGTGGGGAAG)$_2$•$(NH_4)_5$]	-6	2368	1090	[(b)]	X	
Quadruplex dimer [(dGAGGGTGGGGAGGGTGGGGAAG)$_2$•$(NH_4)_5$]	-7	2030	1098	*JACS, **2008**, 130, 10208*	X	
Quadruplex dimer [(dGAGGGTGGGGAGGGTGGGGAAG)$_2$•$(NH_4)_5$]	-8	1779	1113	[(b)]	X	

(a) *JACS = J. Am. Chem. Soc.: JASMS = J. Am. Soc. Mass Spectrom.; IJMS = Int. J. Mass Spectrom.*

(b) *V. Gabelica, E.S. Baker, M.T. Bowers, unpublished results.*

Table S4. Key structural descriptors of the 12-mer and 18-mer triplexes in aqueous solution obtained by averaging the last 50 ns of 100 ns trajectories. For groove nomenclature see Shields *et al*. (Shields, G. C., Laughton, C. A., & Orozco, M. (1997) *J. Am. Chem. Soc.* **119,** 7463). Distances are in Å and angles in degrees. Standard deviations in the average are also displayed.

		$(d(TC^+)\text{-}d(GA)\cdot d(TC))_6$	$(d(TC^+)\text{-}d(GA)\cdot d(TC))_9$
RMSd	all	2.3 +/- 0.4	2.23+/- 0.33
	central	1.5 +/- 0.2	2.05+/- 0.29
% H-bonds (all)	WC	96	96
	Hoogsteen	91	89
% H-bonds (central)	WC	99	100
	Hoogsteen	100	99
Average phase angle	$d(TC^+)_6$	100 +/- 26	110 +/- 44
	$d(GA)_6$	139 +/- 28	139 +/- 28
	$d(TC)_6$	108 +/- 26	116 +/- 45
Groove width	minor	12.89 +/- 1.31	12.13 +/- 1.71
	minor-Major	11.61 +/- 0.65	11.68 +/- 0.53
	Major-Major	20.57 +/- 0.59	20.59 +/- 0.59
Twist	GA	31.34 +/- 3.38	31.57 +/- 3.6
	AG	30.53 +/- 4.15	29.63 +/- 3.63
Roll	GA	1.74 +/- 5.05	1.83 +/- 4.93
	AG	3.73 +/- 5.32	2.77 +/- 5.37
Slide	GA	-1.17 +/- 0.36	-1.22 +/- 0.41
	AG	-1.14 +/- 0.53	-1.22 +/- 0.54

Table S5. Theoretical (MD) and experimental estimates of CCS for the triplexes, at bias = 20 V using neutral and acidic conditions. The experimental errors were calculated from the 95% prediction interval of the respective calibration curves, meaning that the true value has 95% chance of lying in this interval. The theoretical error bars were determined by propagation of standard errors in the different estimates obtained from individual trajectories, charge states and temperatures.

Triplex	Charge	Theoretical CCS (372K)	Experimental CCS (pH = 7)	Experimental CCS (pH = 3)
12-mer	in solution	10.2±0.2		
	-5			9.2 ± 0.2
	-6	10.0±0.3	9.7 ± 1.1	9.6 ± 0.3
	-7		10.2 ± 1.3	10.6 ± 0.4
	-8		11.8 ± 1.5	
18-mer	in solution	14.3±0.5		
	-6			12.9 ± 0.3
	-7	13.2±0.1	13.4 ± 1.3	13.3 ± 0.3
	-8	14.3±0.1	14.2 ± 1.5	14.3 ± 0.4
	-9		15.0 ± 1.8	15.4 ± 0.5
	-10			16.6 ± 0.5

Table S6. Key simulation results obtained from the analysis of 10 different charge states of the 12-mer and 18-mer triplexes in the gas phase at different temperatures and total charge states. In every case we report the average RMSd (from solution structure), collision cross section, radii of gyration and the maintenance of hydrogen bonds. When appropriate, the solution values are also given.

18-mer charge state -7

Replicas T = 300K	Av. RMSd(nm)	Av. Rg (nm)	Av. Ccs(nm²)	Hbond% Watson_Crick	Hoogsteen	Total
Solution	0.23 ± 0.03	1.86 ± 0.03	14.3 ± 0.5	96	89	94
1	0.64 ± 0.02	2.05 ± 0.03	13.6 ± 0.5	88	11	47
2	0.68 ± 0.01	2.03 ± 0.01	13.9 ± 0.5	15	8	13
3	0.67 ± 0.01	2.05 ± 0.01	13.5 ± 0.5	59	27	43
4	0.64 ± 0.02	1.80 ± 0.03	12.9 ± 0.4	91	9	50
5	0.44 ± 0.01	1.78 ± 0.01	13.0 ± 0.4	88	24	56
6	0.57 ± 0.02	2.87 ± 0.02	13.2 ± 0.5	26	9	17
7	0.66 ± 0.01	2.87 ± 0.01	12.9 ± 0.3	78	26	56
8	0.60 ± 0.02	1.76 ± 0.01	12.9 ± 0.4	37	13	25
9	0.54 ± 0.01	1.91 ± 0.02	13.2 ± 0.4	46	6	26
10	0.62 ± 0.02	1.75 ± 0.01	12.7 ± 0.4	53	6	29
Av.	**0.60 ± 0.07**	**1.9 ± 0.1**	**13.2 ± 0.4**	**58**	**14**	**36**

Replicas T = 372K	Av. RMSd(nm)	Av. Rg (nm)	Av. Ccs(nm²)	Hbond% Watson_Crick	Hoogsteen	Total
Solution	0.2 ± 0.1	1.85 ± 0.03	14.2 ± 0.5	95	93	94
1	0.59 ± 0.02	1.77 ± 0.02	12.9 ± 0.4	70	12	41
2	0.65 ± 0.02	1.95 ± 0.02	13.7 ± 0.4	42	8	25
3	0.63 ± 0.02	1.86 ± 0.02	13.5 ± 0.4	58	29	43
4	0.60 ± 0.02	1.82 ± 0.02	12.9 ± 0.4	55	7	31
5	0.61 ± 0.03	1.75 ± 0.02	13.1 ± 0.3	63	20	41
6	0.75 ± 0.01	2.07 ± 0.01	13.7 ± 0.5	40	13	26
7	0.68 ± 0.02	1.81 ± 0.01	13.1 ± 0.3	63	7	35
8	0.76 ± 0.02	2.16 ± 0.01	13.9 ± 0.5	33	3	18
9	0.58 ± 0.02	1.77 ± 0.02	12.6 ± 0.4	34	5	29
10	0.76 ± 0.03	1.83 ± 0.02	13.5 ± 0.4	20	21	20
Av.	**0.66 ± 0.07**	**1.8 ± 0.1**	**13.3 ± 0.4**	**48**	**13**	**30**

Continuation Table S6

18mer charge state=-8

Replicas T = 300K	Av. RMSd(nm)	Av. Rg (nm)	Av. Ccs(nm²)	Hbond% Watson_Crick	Hoogsteen	Total
Solution	0.23 ± 0.03	1.86 ± 0.03	14.3 ± 0.5	96	89	94
1	0.71 ± 0.04	2.10 ± 0.03	14.4 ± 0.5	30	30	30
2	0.64 ± 0.02	2.06 ± 0.03	13.8 ± 0.5	51	42	47
3	0.65 ± 0.03	1.86 ± 0.04	13.7 ± 0.5	42	20	32
4	0.80 ± 0.03	2.22 ± 0.04	14.5 ± 0.4	23	27	25
5	0.73 ± 0.03	1.97 ± 0.03	13.7 ± 0.4	54	23	40
6	0.81 ± 0.02	2.23 ± 0.03	14.4 ± 0.5	39	30	35
7	0.70 ± 0.03	2.09 ± 0.03	14.3 ± 0.5	41	18	30
8	0.70 ± 0.03	2.12 ± 0.03	14.2 ± 0.5	48	16	34
9	0.85 ± 0.03	2.01 ± 0.02	14.1 ± 0.4	35	33	34
10	0.62 ± 0.02	2.82 ± 0.03	14.2 ± 0.5	33	35	34
Av.	**0.72 ± 0.07**	**2.1 ± 0.1**	**14.1 ± 0.5**	**49**	**36**	**43**

Replicas T = 372K	Av. RMSd(nm)	Av. Rg (nm)	Av. Ccs(nm²)	Hbond% Watson_Crick	Hoogsteen	Total
Solution	0.2 ± 0.1	1.85 ± 0.03	14.2 ± 0.5	95	93	94
1	0.76 ± 0.06	1.94 ± 0.03	14.7 ± 0.4	30	30	30
2	0.73 ± 0.02	2.10 ± 0.03	14.1 ± 0.3	51	42	47
3	0.93 ± 0.02	2.12 ± 0.05	15.3 ± 0.5	42	20	32
4	0.82 ± 0.02	2.26 ± 0.04	14.6 ± 0.2	23	27	25
5	0.73 ± 0.03	2.02 ± 0.03	14.3 ± 0.4	54	23	40
6	0.74 ± 0.02	2.04 ± 0.03	13.9 ± 0.4	39	30	35
7	0.70 ± 0.05	1.99 ± 0.06	14.0 ± 0.3	41	18	30
8	0.85 ± 0.03	2.08 ± 0.06	14.6 ± 0.3	48	16	34
9	0.84 ± 0.07	1.97 ± 0.06	14.7 ± 0.4	35	33	34
10	0.84 ± 0.04	2.04 ± 0.05	13.8 ± 0.3	33	35	34
Av.	**0.79 ± 0.07**	**2.0 ± 0.1**	**14.4 ± 0.3**	**49**	**27**	**37**

Continuation Table S6

12-mer charge state -6

Replicas T = 300K	Av. RMSd(nm)	Av. Rg (nm)	Av. Ccs(nm²)	Hbond% Watson_Crick	Hoogsteen	Total
Solution	0.23 ± 0.04	1.36 ± 0.02	10.2 ± 0.2	96	91	94
1	0.63 ± 0.02	1.61 ± 0.02	10.1 ± 0.4	63	37	51
2	0.71 ± 0.01	1.48 ± 0.01	9.8 ± 0.3	56	30	45
3	0.63 ± 0.01	1.43 ± 0.02	9.9 ± 0.3	34	25	31
4	0.65 ± 0.02	1.53 ± 0.01	10.1 ± 0.3	58	30	47
5	0.71 ± 0.01	1.60 ± 0.01	10.2 ± 0.3	48	27	40
6	0.75 ± 0.01	1.56 ± 0.01	10.1 ± 0.4	73	30	48
7	0.76 ± 0.02	1.70 ± 0.02	10.3 ± 0.3	38	17	28
8	0.79 ± 0.01	1.48 ± 0.01	9.9 ± 0.3	45	11	30
9	0.78 ± 0.01	1.50 ± 0.02	9.4 ± 0.3	46	20	35
10	0.53 ± 0.02	1.32 ± 0.01	9.8 ± 0.3	61	37	52
Av.	**0.69 ± 0.08**	**1.52 ± 0.1**	**9.9 ± 0.3**	**52**	**26**	**40**

Replicas T = 372K	Av. RMSd(nm)	Av. Rg (nm)	Av. Ccs(nm²)	Hbond% Watson_Crick	Hoogsteen	Total
Solution						
1	0.60 ± 0.01	1.53 ± 0.06	10.1 ± 0.4	56	24	42
2	0.67 ± 0.04	1.45 ± 0.02	9.8 ± 0.3	16	22	19
3	0.61 ± 0.02	1.47 ± 0.01	9.9 ± 0.4	20	8	25
4	0.84 ± 0.02	1.58 ± 0.07	10.5 ± 0.3	58	12	29
5	0.65 ± 0.05	1.60 ± 0.03	10.2 ± 0.3	44	14	27
6	0.65 ± 0.03	1.58 ± 0.01	10.1 ± 0.4	37	13	25
7	0.74 ± 0.02	1.58 ± 0.02	10.4 ± 0.3	38	17	27
8	0.79 ± 0.02	1.54 ± 0.01	9.9 ± 0.3	23	7	16
9	0.72 ± 0.01	1.37 ± 0.02	9.4 ± 0.4	26	18	22
10	0.30 ± 0.01	1.32 ± 0.01	9.8 ± 0.3	35	21	30
Av.	**0.67 ± 0.06**	**1.52 ± 0.07**	**10.0 ± 0.4**	**35**	**15**	**26**

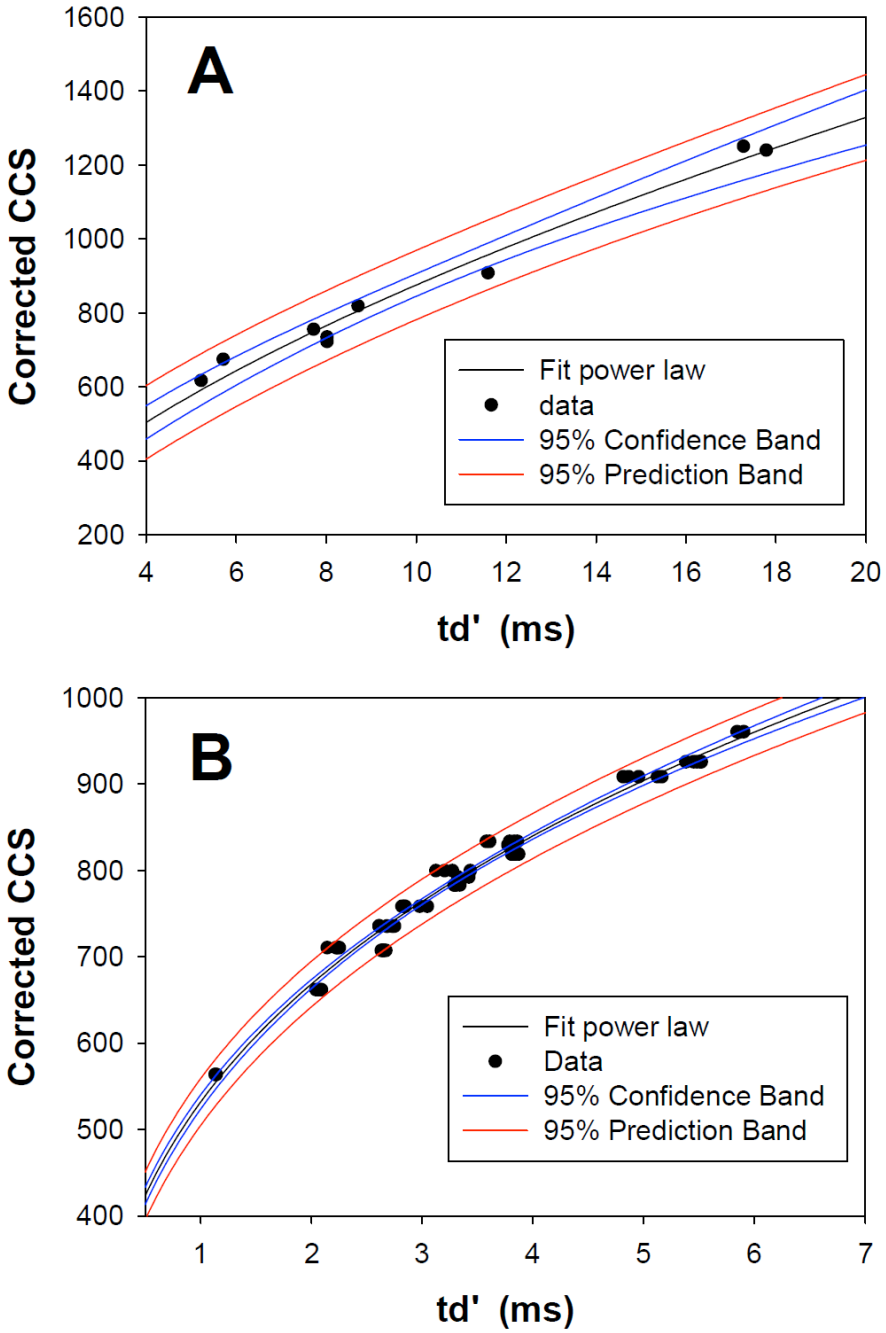

Figure S1: Collision cross section calibration curves: (A) Barcelona; IMS bias = 20 V. (B) Liège; IMS bias = 10, 15, 20, 25 and 30 V. Calibration of the IMS was done with different oligonucleotides in the two instruments. The calibrants lists are given in Table S3. The arrival times were corrected for the traveling time outside the IMS and the traveling time of one wave in the IMS and transfer, yielding t_d' values that are plotted against the corrected collision cross sections (corrected CCS), following the protocol of Ruotolo et al (Ruotolo, B. T.; Benesch, J. L.; Sandercock, A. M.; Hyung, S. J.; Robinson, C. V. *Nat.Protoc.* **2008,** *3*, 1139-1152). A power law was used to fit the data with Sigmaplot 11.0. The 95% prediction interval of the power calibration was used to estimate the error on the experimental CCS. This is therefore the CCS interval in which one has 95% chances of finding the true value using the prediction of the calibration curve.

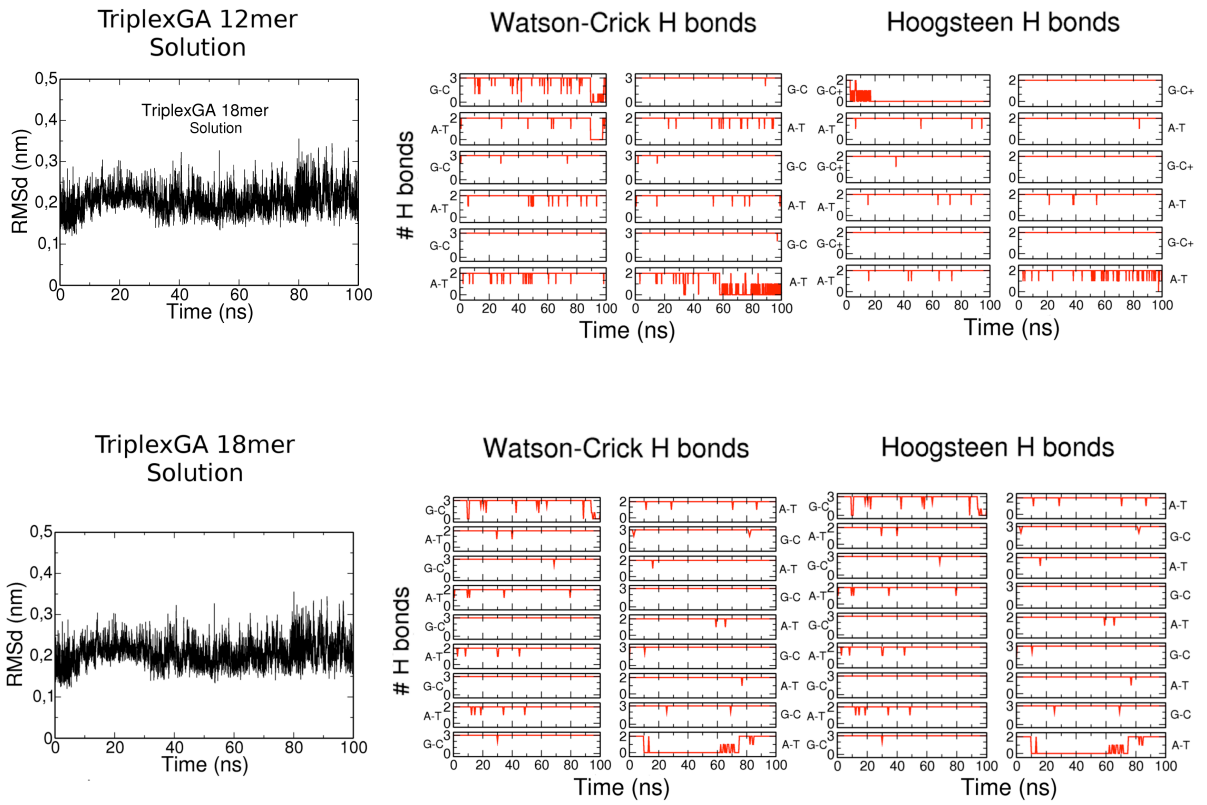

Figure S2. Time evolution of selected structural descriptors of the 12-mer (TOP) and 18-mer (BOTTOM) triplexes in aqueous solution. Left panels, root mean square deviation curves (RMSds) from canonical triplex structure. Middle and Right panels correspond to numbers the Watson Crick and Hoogsteen hydrogen bonds for the different base pairs, respectively.

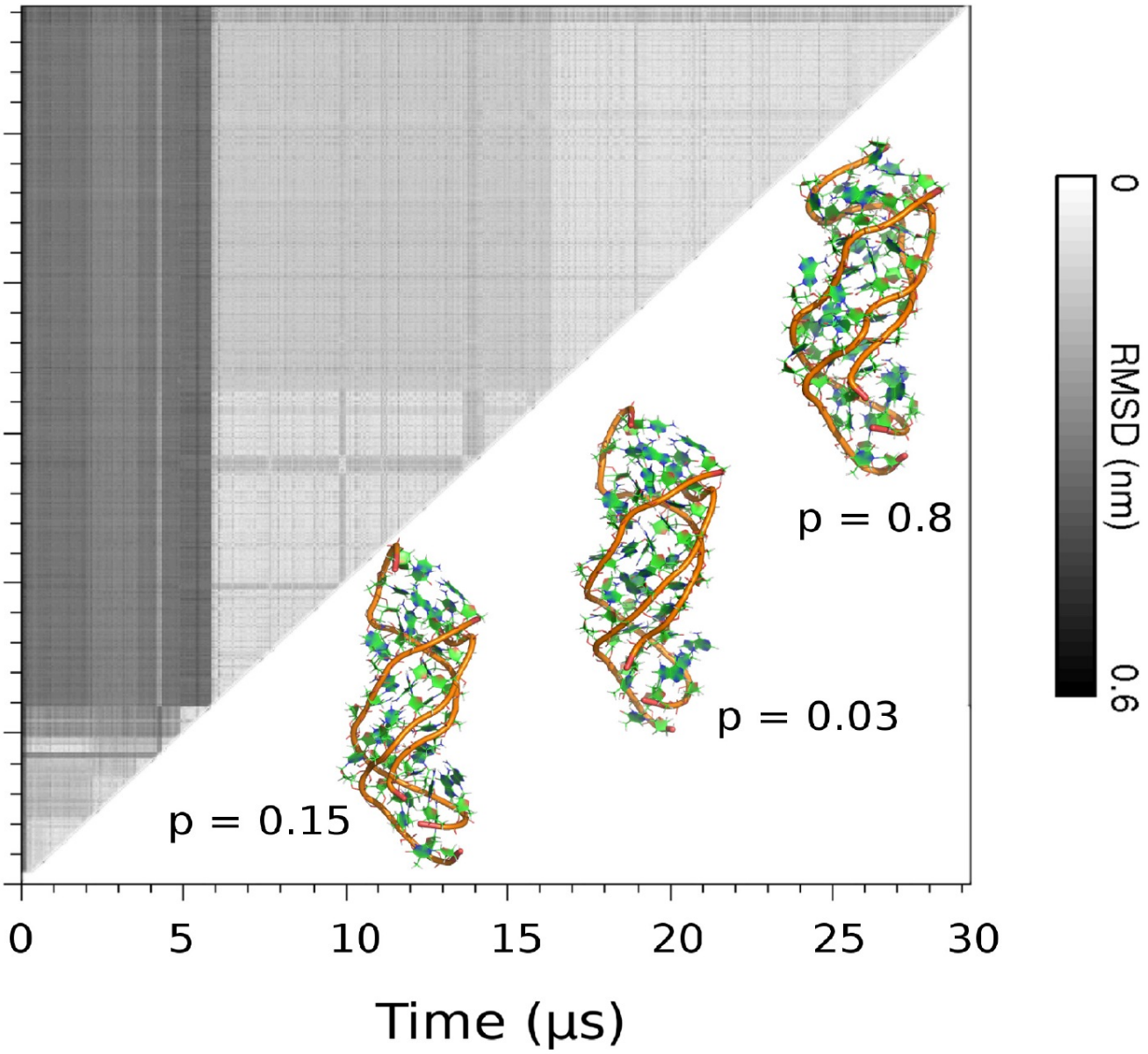

Figure S3. Bidimensional RMSd map for the 30 μs trajectory of the 12-mer triplex in the gas phase. Representative structure of cluster obtained at different simulation frames are displayed, noting the weight of each cluster in the total ensemble.

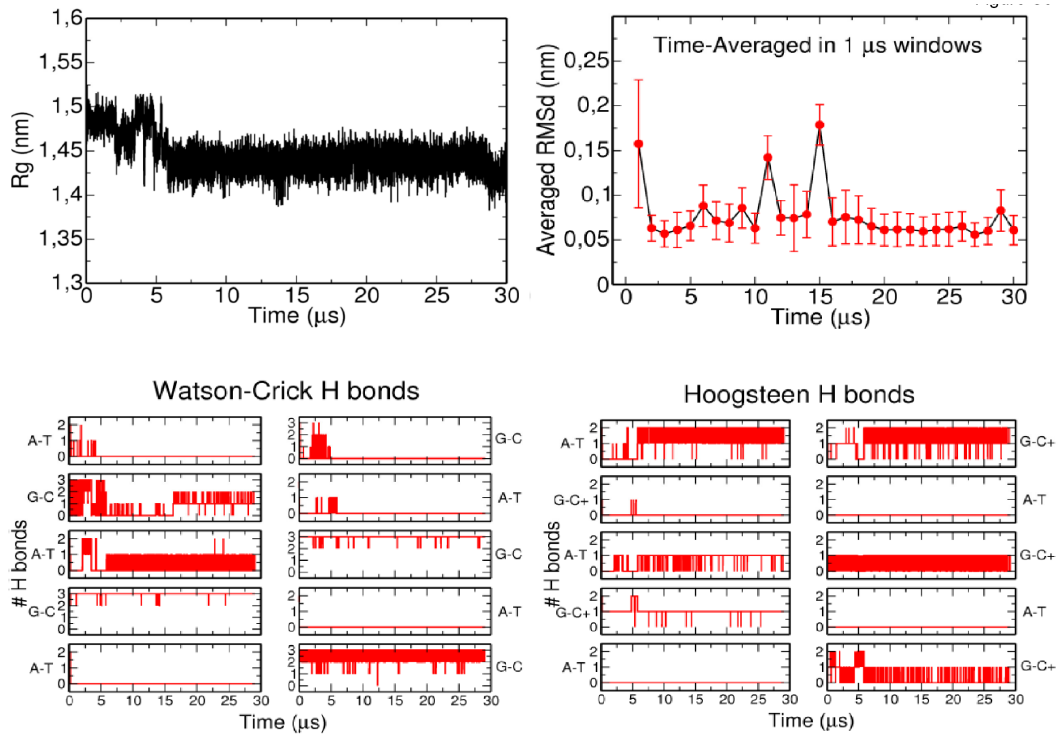

Figure S4. Time variation of several structural descriptors in the 30 μs trajectory of the 12-mer triplex in the gas phase. TOP-LEFT: Radii of Gyration. TOP-RIGHT: average RMSd (with respect to MD-averaged structure) in one microsecond windows. BOTTOM: number of Watson-Crick and Hoogsteen hydrogen bonds.

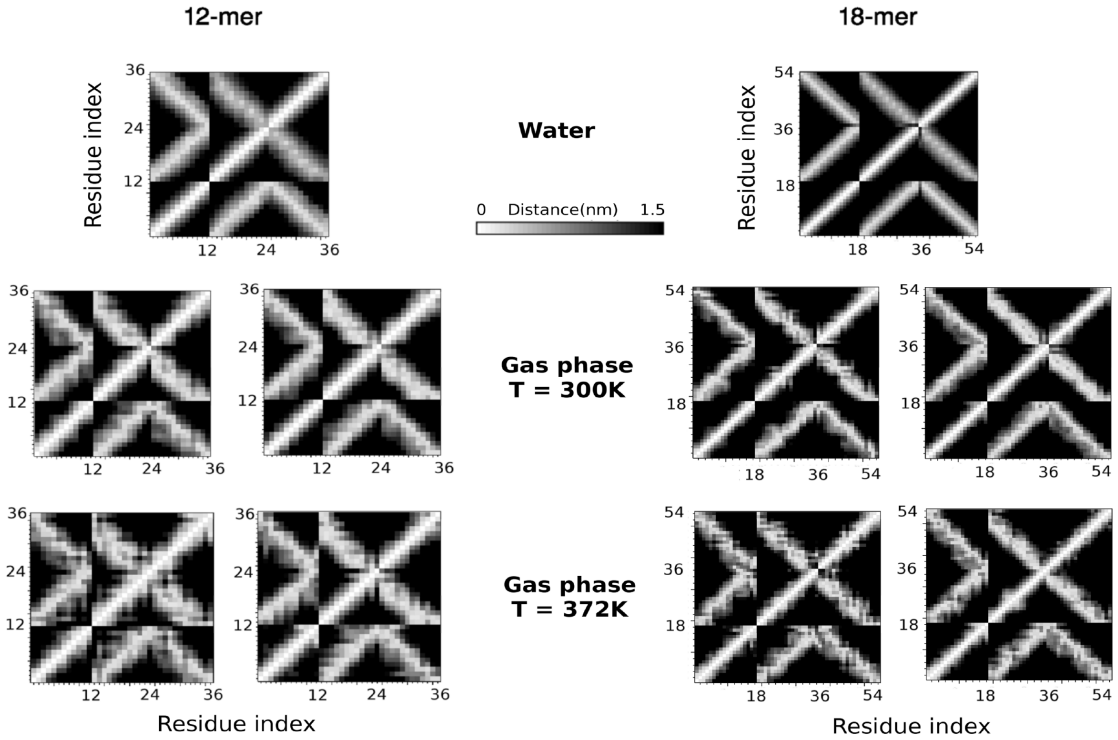

Figure S5. Examples of contact maps found in gas phase simulations (300K and 372K) for 12-mer and 18-mer triplexes. Charge states displayed were randomly selected. The reference maps obtained in aqueous solution are shown as reference.